Reanalysis of Survival of Oscar Winners

TO THE EDITOR: In this issue, Sylvestre and colleagues (1) correctly comment that survival statistics are fallible. The primary analysis in our study (2) was based on the Kaplan–Meier method because life expectancy is the preferred metric in medical decision analysis (3). Our article also provided 40 other secondary analyses to explore different models because no one statistic is ideal. Sylvestre and colleagues argue that the multivariate-adjusted Cox proportional hazards model with a time-varying step function is preferred over our primary analysis approach, do not discuss the limitations of such models, and intimate that other models give an unfair advantage. This position disagrees with us and with other reviews involving our work (4, 5).

We agree that time-varying functions are valuable for addressing a change in status from winning. One drawback with such models can be in assuming the same hazard for all winners following the first win; for example, Jodie Foster (who first won at age 25 years) and Judi Dench (who first won at age 62 years) are assigned identical hazards from age 63 years until death. However, we found that earlier wins were associated with greater advantages, contrary to this assumption. Adding fixed covariates that additionally model age (linear or quadratic) is no simple solution because the likelihood of winning is no simple function of age. The models also have limited assumption. Adding fixed covariates that additionally model age (linear or quadratic) is no simple solution because the likelihood of winning is no simple function of age. The models also have limited power on small data sets, assume no unmeasured heterogeneity, and rarely capture complex trajectories (for example, multiple films, nominations, and wins) (6–9).

We thank many scientists for analyses of our database. We have also done an update to 29 March 2006 and observed 122 more individuals and 144 more deaths since our first publication. Our study (2) was based on the Kaplan–Meier method because life expectancy is the preferred metric in medical decision analysis (3). Our article also provided 40 other secondary analyses to explore different models because no one statistic is ideal. Sylvestre and colleagues argue that the multivariate-adjusted Cox proportional hazards model with a time-varying step function is preferred over our primary analysis approach, do not discuss the limitations of such models, and intimate that other models give an unfair advantage. This position disagrees with us and with other reviews involving our work (4, 5).

We agree that time-varying functions are valuable for addressing a change in status from winning. One drawback with such models can be in assuming the same hazard for all winners following the first win; for example, Jodie Foster (who first won at age 25 years) and Judi Dench (who first won at age 62 years) are assigned identical hazards from age 63 years until death. However, we found that earlier wins were associated with greater advantages, contrary to this assumption. Adding fixed covariates that additionally model age (linear or quadratic) is no simple solution because the likelihood of winning is no simple function of age. The models also have limited power on small data sets, assume no unmeasured heterogeneity, and rarely capture complex trajectories (for example, multiple films, nominations, and wins) (6–9).

We thank many scientists for analyses of our database. We have also done an update to 29 March 2006 and observed 122 more individuals and 144 more deaths since our first publication. Our primary unadjusted analysis shows a smaller survival advantage of 3.6 years (79.7 years vs. 76.1 years; P = 0.005). Applying model 1 of Sylvestre and colleagues’ Appendix so that winners are treated in a time-varying manner yielded a change in mortality of −8% (95% CI, −14% to 26%; P = 0.455). Modifying model 1 so that both winners and nonwinners are treated in a time-varying manner yielded a change in mortality of −15% (CI, −6% to 31%; P = 0.140). These estimates overlap earlier results. Apparently, the survival advantage depends on the analytic method chosen.

The statistical debate concerns built-in survival advantages that yield an immortality bias. We provided methods for addressing this bias, observed multiple findings suggesting this bias was not large in our cohort, and estimated the hidden confounding that would need to be postulated. We found no survival advantage when we compared individuals with many nominations and individuals with no nominations, for example, contrary to estimates of a large immortality bias. Moreover, we presumed that individuals not reported dead were alive, which is a different type of immortality bias that causes almost all of our analyses and Sylvestre and colleagues’ analyses to underestimate survival differences.

Donald A. Redelmeier, MD
Sheldon M. Singh, MD
University of Toronto
Toronto, Ontario M4N 3M5, Canada

Potential Financial Conflicts of Interest: None disclosed.

References

EDITORS’ NOTE: The debate between Sylvestre and colleagues (1) and Redelmeier and Singh shows both the value and limitations of prepublishing peer review and underscores the importance of review after publication. The original paper by Redelmeier on the survival of Oscar winners (2) underwent close in-house scrutiny and external methodologic review, which resulted in several new analyses, including the “time-varying covariate” model we discuss here. Because the editors felt that the methodologic issues were subtle, we also took what was at that time a somewhat unusual step to facilitate postpublication review. As a condition of publication, we required the authors to make the data set available to interested researchers. Unfortunately, various complications prevented its prompt dissemination, and it has taken almost 5 years for someone to come forward with a reanalysis of the data. We are glad to publish Sylvestre and colleagues’ reanalysis, partly because the article affords a chance to amend a widely publicized result, but more so because the analytic methods at issue apply to many health care research questions.

The main purpose of this letter is to help the technically less sophisticated reader to understand the issues under discussion. The central issue is how best to analyze a sudden change in risk due to some life event (becoming ill, starting a high-risk behavior, or starting a treatment). In this case, the event is a salutary one: winning an important prize. The question is exactly when to “start the clock” in assessing whether the prize changes the winner’s subsequent risk profile, and how to do that analytically. Redelmeier and Singh referred to this question in their original paper as the “time-zero” problem. Because Redelmeier and Singh matched winners and nonwinners on their age at the time the Oscar was won, their analysis appeared to start the clock at the right moment. However, their primary analysis did not maintain that matching; instead, it combined all winners into one group and all losers into another group and compared winners’ and nonwinners’ survival from birth. With this approach, winning the prize gets credit for how long the winner lived before winning the prize. This primary analysis produced a large and highly statistically significant advantage (a 3.9-year increase in life expectancy, equivalent to a 28% annual risk reduction), the outcome high-
lighted in the original paper and abstract and publicized in subsequent media reports.

As Sylvestre and colleagues make clear, the optimal methods of analysis involve starting the clock at the moment of winning the prize. In their 2001 paper, Redelmeier and Singh presented a number of secondary analyses that started the clock at different moments, including a Cox survival analysis in which the risk for subsequent death for winners and nonwinners could change at the instant of winning an Oscar. With this form of analysis, the putative risk modifier—in this case, winning the prize—would have no effect early in a prizewinner’s life but would have an effect after the win. Winning the prize is, in statistical terminology, a “time-varying covariate.” The Cox model suggested a 20% mortality risk reduction, with borderline statistical significance, and a range of uncertainty that just included the possibility of no survival benefit. Speaking for the Annals Editors, we regret that the original paper did not adequately emphasize this more equivocal but probably more correct result.

In the preceding letter, Redelmeier and Singh report the results of using the time-varying covariate modeling approach to analyze their most recently compiled data set of Oscar winners (updated to 2006). This analysis yields still weaker, now statistically nonsignificant evidence that winning an Oscar prolongs life: either an 8% survival advantage (with statistically compatible effects ranging from as low as 14% shorter survival to as high as 26% longer survival) or a 15% survival advantage (the uncertainty of which is compatible with a range of 6% shorter survival to up to 31% greater survival). The 2 estimates differ according to how the analysis handles the nonwinners.

Sylvestre and colleagues point out that although this Cox “time-varying” result is closer to the truth than the result that Redelmeier and Singh reported as the primary analysis in their paper, it may not yet be optimal, for many of the same reasons that Redelmeier and Singh point out in their letter. Sylvestre and colleagues prefer the conceptually simpler approach of measuring life expectancy from the moment of winning the Oscar. This approach, outlined in their Web-only appendix, produces a result qualitatively consistent with the result from the time-varying model that Redelmeier and Singh report in their letter.

The debate about whether winning an Academy Award confers any survival advantage—and if it does, by how much—will continue in exchanges between interested scientists. To facilitate their participation in this discussion, we are posting on the Annals Web site the data set (updated to March 2006) that Redelmeier and Singh have provided and that Sylvestre and colleagues used in their analysis. The Editors invite people who want to contribute to the discussion to communicate their ideas as a Rapid Response letter about Sylvestre and colleagues’ article. We hope that other members of the statistical community will take up the challenge of determining the most appropriate way to measure the effect of winning an Oscar and the statistical uncertainty around the result. Their efforts will inform the analysis of many similar phenomena in biomedicine.

When the dust settles, we expect that the estimated effect will be nonsignificant, and closer to Redelmeier and Singh’s adjusted estimates and to the estimate of Sylvestre and colleagues than to the original estimate of 3.9 years (now 3.6 years, using the 2006 updated data set). Until then, we urge everyone to observe much greater caution about claiming the existence of an “Oscar effect” on life span. Granted, doing so may mean some tempering of joy among Academy Award winners. They will get their statuette, and the attention it brings, but we doubt that winning it will confer many—if any—more years to enjoy the fruits of their enhanced celebrity.

Steven Goodman, MD, PhD
Associate Editor

Harold C. Sox, MD
Editor

Potential Financial Conflicts of Interest: None disclosed.

References

Cryptogenic Stroke and Patent Foramen Ovale

TO THE EDITOR: In their comprehensive and informative Update (1), Drs. Holloway and Józefowicz suggest using warfarin for secondary prevention of stroke in patients with atrial septal defect. The current literature has no strong evidence to support this view, and therefore the current guidelines from the American Academy of Neurology state that “the evidence is insufficient to determine whether aspirin or warfarin is superior in preventing recurrent stroke or death in patients with patent foramen ovale (PFO) alone” (2). However, the American Academy of Neurology does recommend warfarin therapy in patients with patent foramen ovale and evidence of deep venous thrombosis (2).

The rationale for aspirin therapy in patients with patent foramen ovale comes from a French study of 216 patients with a cryptogenic stroke (3). This trial reported that the incidence of recurrent stroke was only 2.3% after 4 years in patients who had patent foramen ovale alone and were taking aspirin, a value similar to the 4.2% risk in the control group. Support for the use of aspirin also comes from the Patent Foramen Ovale in Cryptogenic Stroke Study, which did not demonstrate a statistical difference between the effects of aspirin and warfarin on the risk for subsequent stroke or death among patients with cryptogenic stroke and patent foramen ovale (4). Although studies have favored warfarin over aspirin for secondary prevention of stroke in patients with patent foramen ovale and atrial septal defect, they included small numbers of patients, had limited statistical power, and were unblinded and retrospective (5). On the basis of currently available evidence, the American College of Chest Physicians also recommends aspirin over no therapy or warfarin therapy in patients with patent foramen ovale (6).

Ashok K. Malani, MD
Husam Ammar, MD
Heartland Regional Medical Center
St. Joseph, MO 64506

Potential Financial Conflicts of Interest: None disclosed.
Letters

References

IN RESPONSE: In our article, we do state that “warfarin therapy is generally not indicated for secondary stroke prevention except for patients with transient ischemic attack or stroke in the setting of persistent or paroxysmal atrial fibrillation and for some patients with a documented hypercoagulable state, left ventricular ejection fraction of 0.3 or less, carotid or vertebral artery dissection, or patent foramen ovale with an atrial septal defect.” We do not make this recommendation for patients with only a patent foramen ovale.

The American Academy of Neurology guideline states, “It is possible that the combination of a patent foramen ovale and atrial septal aneurysm confers an increased risk of subsequent stroke in medically treated patients who are less than 55 years of age” (1). The guideline concludes that there is insufficient evidence to determine the superiority of aspirin or warfarin for the prevention of stroke or death in this situation.

Given the increased risk and the lack of evidence to guide proper therapy, we do still conclude that warfarin may be considered in some patients with transient ischemic attack or stroke who have both a patent foramen ovale and an atrial septal aneurysm (for example, younger patients). We did not mean to imply that warfarin was “indicated,” but that it may be preferred and used after thoughtful estimations of the benefits and burdens of each therapy and incorporation of the preferences of the patient in terms of tolerance and acceptance of treatment and outcome risk.

Robert G. Holloway Jr., MD, MPH
Ralph F. Józefowicz, MD
University of Rochester School of Medicine
Rochester, NY 14642

Potential Financial Conflicts of Interest: None disclosed.

Reference

Poorly Controlled Cardiovascular Risk Factors and ICD-9-CM Codes

TO THE EDITOR: The article by Rodondi and colleagues (1) on therapy modification in response to poorly controlled cardiovascular risk factors ignores the reality that the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) classifies patients with “poorly controlled” cardiovascular risk factors as “well controlled.” Accordingly, physicians using such terms as poorly controlled to describe their uncontrolled, diabetic, hypertensive, and hyperlipidemic patients must have their records coded as “well controlled,” since ICD-9-CM requires strict interpretation of physician diagnosis assignment. Similarly, patients with “hypertensive emergency” and “hypertensive crisis” will be coded as having well-controlled hypertension by ICD-9-CM unless the physician uses the official Centers for Disease Control and Prevention–sanctioned terms accelerated or malignant to describe their hypertension.

As a result, ICD-9-CM administrative databases measuring physician quality and resource consumption (for example, Healthgrades, Solucient, APR-DRGs [All Patient Refined Diagnosis-Related Groups], and Medicare’s proposed Consolidated Severity-Adjusted Diagnosis-Related Groups) misrepresent patient severity of illness and outcomes when physicians use these and other imprecise ICD-9-CM terms to describe their patients’ illnesses. Annals of Internal Medicine and other highly regarded peer-reviewed journals should consider official disease nosology and classifications when editing authors’ manuscripts. Alternatively, the American College of Physicians’ appointed member of the Editorial Advisory Board of the Coding Clinic for ICD-9-CM may advocate for revisions of ICD-9-CM that better match physicians’ vocabularies in the literature and day-to-day practice.

James S. Kennedy, MD
FTI Cambio Health Solutions
Brentwood, TN 37027

Potential Financial Conflicts of Interest: None disclosed.

Reference

IN RESPONSE: We agree that it is difficult to accurately capture disease severity and control with commonly used ICD-9 classifications. For this reason, we used ICD-9-CM codes, along with ambulatory blood pressure measurements, laboratory results, and prescriptions, to identify the presence of hypertension, dyslipidemia, and diabetes mellitus but not to grade their degree of severity or control. As described in our Appendix Table 2 and in our Methods section, we used actual ambulatory blood pressure measurements and laboratory results from the electronic records at Kaiser Permanente to define control and pharmacy records to identify medication intensification. Previous studies have documented the accuracy of the Kaiser Permanente clinical databases used in our study (1, 2). For example, diabetes diagnosis, myocardial infarction, and stroke were all confirmed at chart review in 98%, 99%, and 75% of cases, respectively, as described in our article. Although we cannot exclude some
misclassifications in the identification of hypertension, dyslipidemia, and diabetes mellitus, our diagnostic criteria are certainly more accurate than relying on ICD-9 codes alone.

In our study, levels of control were determined by using actual measurements and current clinical guidelines. Because we found that measuring therapy modifications in response to poor control in a large population was feasible, future studies should examine whether giving physicians feedback on this process-of-care measure may increase levels of control. This kind of measurement also has limitations but may provide a more accurate index of the quality of clinical care than relying solely on measures examining the proportion of patients whose condition is under control.

Nicolas Rodondi, MD, MAS
University Outpatient Clinic, University of Lausanne
1011 Lausanne, Switzerland

Eve A. Kerr, MD, MPH
Ann Arbor Veterans Affairs Center for Practice Management and Outcomes Research and University of Michigan Medical School
Ann Arbor, MI 48113

Joe V. Selby, MD, MPH
Kaiser Permanente
Oakland, CA 94612

Potential Financial Conflicts of Interest: None disclosed.

References

Corporate Strategies for Computerization

TO THE EDITOR: McDonald’s critique of computerization (1) highlights advice that all physicians would do well to heed. The ease of bar coding can allow us to become automatons, shedding our critical thinking in the process. Although McDonald touched on technological alternatives and systematic changes that would improve such innovations as bar coding, we suggest some additional lessons that medicine can learn from the corporate world.

Corporate principles can enhance patient safety while maintaining a culture of efficiency and effectiveness. The emergence of Six Sigma principles (2) and quality management departments as applied to medicine is promising. Kiosks have been used extensively at airports and financial institutions. The advent of a patient-centric kiosk, where patients input their information, has potentially great implications for efficiency and patient safety (3). Another common retail practice involves the repetition of an order. With the exception of blood product administration, this repetition is seldom used in medicine. To review a process with a patient sounds so simple, yet it is so difficult to enact in clinical practice. Medicine could learn from business by involving its customers—patients—more.

In addition to the technologies mentioned by McDonald, another that deserves mention is natural language processing (NLP). Still in its infancy, NLP can enhance adverse event detection (4). This technology uses computer algorithms to detect potential adverse events within the confines of an electronic medical record. The key challenge will be to integrate NLP into processes that inform providers in a timely and appropriate fashion, to prevent adverse events. While bar coding can dull our sensibilities, NLP can heighten our senses to adverse events.

It is inevitable that new technologies will emerge as we shuttle patients through a hospital. How we implement these new technologies should not rest solely with a consultant, a vendor, or a committee. Our technology solutions should be looked on in the same way that we care for patients: from head to toe.

Edward C. Wu, MD, MBA
Nirav Shah, MD, MPH
New York University School of Medicine
New York, New York 10010

Potential Financial Conflicts of Interest: None disclosed.

References

IN RESPONSE: I thank Drs. Wu and Shah for their thoughtful letter. Just to be clear, I was not criticizing bar-code systems in general. Bar coding is the perfect solution for many identification problems, for example, the identification of groceries at the checkout counter; we all appreciate the resultant faster checkout times. My criticism was limited to bar-coding systems used to verify medication dispensing at the bedside. Both published reports and recent conversations with disappointed CEOs suggest that these systems have not delivered the nursing time savings and error elimination they promised. Furthermore, they appear to change nursing priorities—an unintended consequence. I am sure that with time and improved technology these systems will meet their mark, but the available evidence contradicts the frothy hype for rushing to implement these systems today.

There is a general lesson here. The health care industry expects information technology, such as bar-coded medication dispensing and physician order entry, to solve many or most of its problems. These are plausible expectations, and many will eventually be realized. However, today, the hype and the hope have run far ahead of the evidence and experience. We need to critically examine the benefits and harms of these systems as implemented in operating environments to determine what works and what doesn’t and to identify design flaws and misassumptions as Patterson and colleagues (1) did for bar coding and others have done for computerized physician order entry. Only with such knowledge will these products improve and reach their full promise.
The expectations about Six Sigma’s potential for health care may also be a bit inflated. Even in manufacturing, the real achievement may only be four and a half, rather than Six, Sigma, and some question the applicability of these principles to fields other than manufacturing (2). Given that health outcomes are not influenced by many-fold increases in care process investment (3) and the fact that care costs have reached crisis levels, one might wonder whether this is the time to eliminate costly but marginal processes rather than investing even more to perfect them.

I agree that methods for capturing data derived from patients have value and am a fan of the work by Hripcsak and colleagues (4). Natural language processing should become a major asset to medical information management.

Clement J. McDonald, MD
Regenstrief Institute
Indianapolis, IN 46202

Potential Financial Conflicts of Interest: None disclosed.

References

Correction

Correction: A Rare Cause of Nonalcoholic Fatty Liver Disease
In a recent letter on nonalcoholic fatty liver disease (1), the order of the authors was incorrect. Judith Fischer, MD, PhD, should have been listed as the third author. The senior author, Christian P. Strassburg, MD, should have been listed last.

Reference